1. По какому правилу байесовский классификатор относит объект к некоторому классу?

1 point

- [x] $a(x) = \operatorname{argmax}_y P(y)P(x \mid y)$
- [ ] $a(x) = \operatorname{argmax}_s L(s, x)$
- [ ] $a(x) = \operatorname{argmax}_y P(y)P(y \mid x)$
- [ ] $a(x) = \operatorname{argmin}_s L(s, x)$

2. В чем заключается «наивная» гипотеза в наивном байесовском классификаторе?

1 point

- [ ] В том, что распределения классов можно оценить по обучающей выборке.
- [ ] В том, что вместо максимизации $P(y \mid x)$ можно максимизировать $P(y)P(x \mid y)$.
- [x] В том, что вместо оценки многомерной плотности $P(x \mid y)$ достаточно оценивать плотность каждого отдельного признака $x^{(j)}$ и положить $P(x \mid y) = P(x^{(1)} \mid y) \cdot \dots \cdot P(x^{(n)} \mid y)$.

3. Как можно оценить по обучающей выборке априорную вероятность класса $P(y)$, если количество объектов в обучающей выборке $\ell$, из них к классу $y$ относятся $l_y$?

1 point

- [ ] $P(y) = \frac{l_y}{l}$
- [ ] $P(y) = \frac{l}{l_y}$
- [ ] $P(y) = \frac{l - l_y}{l}$
- [x] $P(y) = \frac{1}{\frac{l}{l_y} - 1}$

4. Допустим, мы хотим использовать наивный байесовский классификатор, в котором плотности распределений каждого признака (среди объектов заданного класса) восстанавливаются с помощью параметрического подхода, а распределения предполагаются нормальными. Априорные вероятности нам уже известны. Какие еще распределения и как будут восстанавливаться в таком подходе?

1 point

- [ ] $P(x \mid y)$ с помощью построения гистограмм
- [ ] $P(y \mid x)$ с помощью построения и сглаживания гистограмм.

- $P(x)$ и $P(y \mid x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{x-\mu^2}{2\sigma^2}}$, параметры $\mu$ и $\sigma$ свои для каждого класса $y$ и определяются по обучающей выборке с помощью некоторых оценок (например, оценок, полученных по методу максимального правдоподобия).

- (selected) Все так, как в предыдущем варианте, только оценивается не $P(y \mid x)$, а $P(x \mid y)$.
  И $P(x)$ оценивать не надо.

5. Решается задача классификации, метки классов $Y = \{0, 1\}$. Какая функция потерь из предложенных приведет к тому, что $a(x)$ будет оценивать вероятность $P(y \mid x)$?

1 point

- $L(y, a(x)) = |y - a(x)|$
- $L(y, a(x)) = \exp(-ya(x))$
- (selected) $L(y, a(x)) = y \ln a(x) + (1-y) \ln \big(1 - a(x)\big)$